DOCUMENT RESUME

ED 040 449 CG 005 470

AUTHOR Boyd, Robert E.
TITLE Use of the Hill Interaction Matrix in Individual Counseling Research.
INSTITUTION American Personnel and Guidance Association, Washington, D.C.; Illinois Univ., Urbana.
PUB DATE Mar 70
NOTE 14p.; Paper presented at the American Personnel and Guidance Association Convention, New Orleans, Louisiana, March 22-26, 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$0.80
DESCRIPTORS Counseling, *Counseling Effectiveness, Counseling Goals, Counselor Educators, Counselor Evaluation, *Counselor Performance, *Counselor Training, *Research, Research Problems, Research Utilization

ABSTRACT

Research in counseling effectiveness is moving from the use of gross outcome measures to analysis of the counseling process. One successful approach has been the adoption of a social-psychological model of interview analysis which attempts to specify within-interview conditions which facilitate client behavior change. The criteria and procedure for developing such a system is briefly presented and the fidelity of the Hill Interaction Matrix to them elaborated. The paper attempts to provide four types of data about the Matrix: (1) that relative to its measurement characteristics; (2) the uses which have been made of it in individual counseling research, specifically the studies of Lee, Helervik, and Boyd; (3) problems involving its use, primary among which was the training of raters and all of which concerned aspects of the rating system; and (4) suggested extensions of the current scoring procedures. The conclusion holds that the Hill Interaction Matrix fulfills the conditions necessary for instruments of its type better than any other scale currently available. (TL)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED 040449

Use of the Hill Interaction Matrix in Individual Counseling Research

Robert E. Boyd
University of Illinois

Foreword

Author's Note: The following is a brief introduction to the Hill Interaction Matrix' system of statement classification. It must be remembered that the body of this paper was presented as one of a series on the topic and was written for an audience which had previously been given background regarding the scale. Figure 1 is a diagram of the matrix including cell designations and cell weights. The cell weights system was developed by Hill to indicate the hypothesized therapeutic value of statements meeting the criterion for inclusion in that cell. Readers interested in the scale should peruse Hill's 1965 publication entitled Hill Interaction Matrix published by the University of Southern California.

From its inception the Hill scale has been visualized in the form of a matrix. The current scale has two dimensions, one dealing with level and style of content and the other dealing with level and style of therapeutic work. The current form of the scale yields a matrix of 20 cells. In current practice, the top four cells are seldom used with groups other than severely disturbed hospitalized patients.

As can be seen from the accompanying illustration, the content-style categories fall into two areas: non-member centered and member centered. This reflects the type of communication occurring within the confines of the group. Non-member centered communication was divided by Hill into the two specific categories, topic and group. The topic category was defined as any conversation occurring within the group which deals with a subject other than persons or relationships within the group. It covers any subject of general interest to members in the group. Examples are people outside the group, weather, or current events. Category two, group, includes all conversational items which involve discussion within the group about

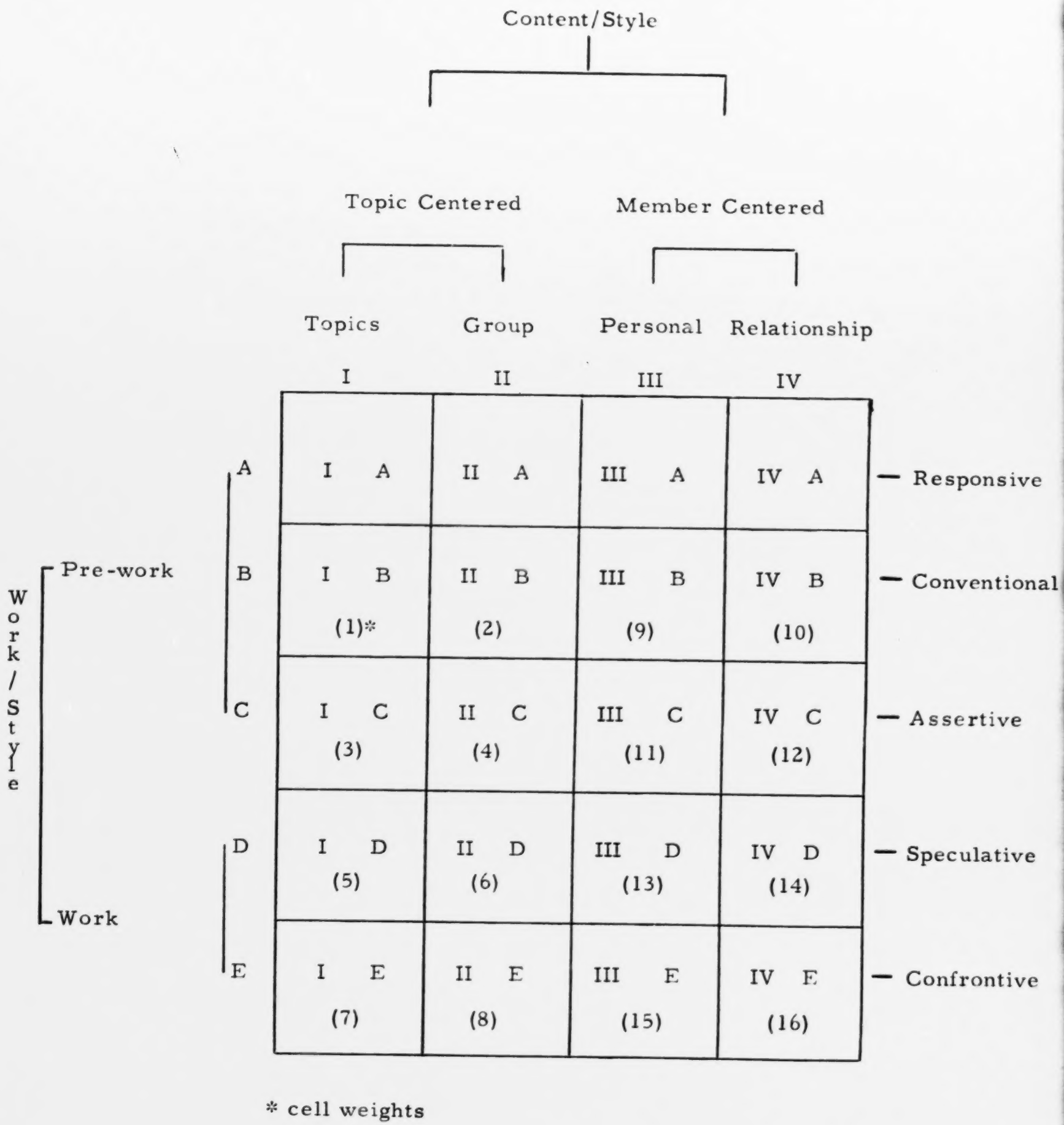
CG 005 470

roles, formation, and general group maintenance. The member centered categories include all conversations of a personal relationship nature. Personal items are responses dealing with the personal actions or feelings of group members by either the topic person or other group members. The relationship category deals with the verbal interaction which gives evidence of a relationship between various members of the group.

Work-style categories are listed along the left hand side of the matrix. These include categories in two areas, pre-work and work. The pre-work area is conceived of as being less productive of personal growth and change than the work areas. Reading from top to bottom the responsive category may be defined as including monosyllable communication, not particularly adding to the ongoing activity of the group but intimating some very slight level of involvement in the group activities. The conventional category includes statements regarding facts and information about the interview content. The information is generally appropriate and there is no particular problem involved in the gathering of data. The assertive category typically deals with hostile, attacking, definitive statements which shut off discussion rather than encouraging it and hence limit the opportunity for personal growth. Work areas include the speculative and confrontive categories. Speculative statements may be defined as statements open to two-way conversation in which one person invites the other to examine the issues which have been presented. This is a high risk area which is tempered with statements such as, "I think," I believe," and "it's possible," allowing the individual who makes the statements a graceful way of escaping from his opinion, yet causing the recipient of the statement to view his previous comments in light of the somewhat threatening state-

Figure 1

Hill Interaction Matrix



* cell weights

ment that may allow more personal growth to ensue. Confrontive statements vary from speculative statements in that the risk level is higher and the confronter is more willing to state in definitive terms his opinions. He often provides some form of clarification or documentation to support his statement. A typical speculative statement such as "I think" changes to "what you really need to do;" specific comments such as this involve more personal risk-taking and a higher personal investment.[2]

Paper presented at the American Personnel and Guidance Association Convention, New Orleans, 1970.

There is not much doubt that research in counseling effectiveness is moving from using gross outcomes measures to an analysis of the counseling process in an attempt to specify within-interview conditions which facilitate client behavior change.[18] A look at the professional literature of the past decade abounds with attempts to specify the counselor interview behaviors necessary for such change; number of responses per interview, talk ratio, physiological changes, and hypothesized constructs such as genuineness have all been used in attempts to specify interview behavior helpful to the client.

One of the more successful approaches used by researchers has been to adopt a social-psychological model of interview analysis. This necessitated the classification of within-interview verbal behavior into descriptive categories having some internal consistency and being mutually exclusive in character. A system of this nature is necessary if units of behavior are to be defined in such a manner that they may be reliably observed and form a valid categorization system.[18]

Such a system should be defined precisely enough to permit scaling of observational units on a continuum. This allows the utilization of units weighted on the basis of therapeutic impact and enhances the instrument's use as it facilitates investigation of specific aspects of interview behavior. The units so established must reflect operational definitions of the criteria sought.[18] The Hill Interaction Matrix approximates these conditions. At variance with

the Truax and Carkhuff Scales[16] and the Burks Scales,[3] the HIM was developed from an analysis of the therapeutic process, and was not the product of a particular theory or theories of counseling.

The HIM specifies classes of verbal behavior indicative of the interview communications process. This classification system was developed based upon the actual behavior occurring within the interview, thus allowing the researcher to establish a functional relationship between 1) the interview behavior and the scale, and 2) the interview behavior and resultant post-counseling behavior.[12]

Early work in interactional analysis was done by Bales, basically along three dimensions; getting information, making decisions, and carrying out actions.[14] At about the same time Timothy Leary and his associates developed an interview analysis system which utilized a circumplex having descriptive categories which called upon the rater to make decisions relative to the mood of the themes being discussed.[5]

Hill began to develop his instrument about 1954. The development of its structure and form was based upon his experiences as a group therapist.[8] Later, Truax and Carkhuff published their initial research based upon an analysis of interview data in individual counseling.[16] In 1968 Burks developed a series of five scales for use in interview analysis.[3] One of his scales was unique in that it recorded non-verbal behavior and related it to counseling process.

In an attempt to satisfy specificity of observed behavior Hill's first scale contained 108 cells. While his first scale increased preciseness of categories,

it proved too complex as a system. His further work resulted in the current 20 cell matrix which provided a readily usable instrument. He then developed cell definitions precise enough to provide the user with a clear understanding of which specified verbal behaviors were to be included in each cell. That Hill satisfied this criterion is supported by his report that interrater reliability coefficients of .78 and higher are readily obtainable using his matrix with therapy groups.[8] Boyd reported that when using the HIM in dyad analysis he has generated interrater reliability figures in the .90's.[2] It must be remembered that dyad analysis is much less complex than the usual group analysis. Hill's final scale, then, is a compromise between specificity of behavior and the number of categories raters can handle reliably, given reliable category decision rules.

Validity is also an issue. Hill reports general predictive validity for his therapy groups which indicates that groups expressing larger numbers of behaviors recorded in his high weight cells tend to discharge more group members and receive more "I feel better" statements. The repeated replicability of the usefulness and similarity of findings by various researchers would seem to lend validity to the scale. One recently reported study by Seligman and Sterne, while a process study, did indicate results in the predicted direction utilizing the HIM as the criterion measure.[13] There would, therefore, seem to be reason to grant that the scale possesses enough validity to be useful, given that the raters are trained well enough to score interview statements in a reliable manner.

The training of raters is an issue deserving of some discussion. Boyd has reported good results with raters having as little as four hours of training.[2] Anderson, in personal conversation, has stated it takes 150 hours to train a good rater.[1] Hill suggests the use of the Mark I and Mark II to train raters. These are decks of cards, each with a single statement which has been pre-rated against the matrix to provide potential raters with actual standard experiences in rating. Another training approach involves the rating of interview flow with a "valid" rater until the desired degree of reliability results. Parenthetically, the concept of a "valid" rater was discussed by Cannon. His point was that it is possible for raters to produce high interrater reliability, but to be reliably inaccurate![4] Both approaches, and variations thereof, assume that the raters have a good understanding of the system and standard definitions of each category in mind.

Why such a training time differentiation? A search of the recent literature seems to support the hypothesis that the richer the psychological and experiencial background of the raters, the faster they grasp the system and the more reliably they rate. Cannon's research supports this position.[4] The work of Vingoe and Antonoff also tends to be supportive.[17] This hypothesis was generated by the difference in rater training problems experienced by Lee and Hellervik in joint dissertations at the University of Minnesota using undergraduates in psychology as raters and Boyd's experience using raters who held at least the master's degree in counseling or psychology and who had experience working in the field as a professional.[2, 7, 11]

Obviously, the specification of the instrument's parameters, both in terms of instrumentation and the competencies necessary for raters, is of prime importance if the matrix is to generate replicable results. The Hill Interaction Matrix, therefore, would appear to satisfy Peterson's four requirements for a useful research tool:

1. the parameters are relatively clear cut
2. the categories are inclusive for homogenity and exclusive of all other classes of behavior
3. it has proven useful in diverse settings, and
4. it is easily teachable. [12]

The HIM is a numerical category scale. The continuum upon which it is based is the hypothesized therapeutic value of each interview statement. While such hypothesized values may appear arbitrary in nature, they are the results of many hours of interview analysis by practitioners in the field, primarily Hill and his associates and, hence, are the result of professional decisions by practitioners. The resultant cell weight system did coincide with his outcomes research on group members. lending credence to the assigned cell values.

Three known studies have utilized the HIM as a criterion measure in individual counseling research. Lee and Hellervik adopted the scale to dyad research by the simple expediency of changing the group category to dyad. [7, 11] The simplicity of the change necessary indicates how generally usable the scale is in interaction analysis. In their studies they used "target" areas of the matrix in a behavior modification experiment. Lee used as his target the

three high weight cells in the lower right hand corner of the matrix-relationship-speculative, relationship-confrontive, and personal-confrontive. His goal was to train counselors to approximate statements of these types. Hellervik used two target areas, the first, for his experimental group, was comprised of the same three cells as Lee's, the second the low weight matrix area in the upper left hand corner of the matrix. Cells weighted one through five were the control group target. He hypothesized that the control group would be less therapeutic than the reinforced experimental group. Experimental results were reported in terms of learning curves descriptive of the interactive behavior which occurred in each interview.

The third study, Boyd's, used the entire HIM matrix.[2] It hypothesized that certain personality variables affected within-interview behavior and hence outcomes. Results were reported using the more traditional analysis of variance procedure.

As it is possible to generate equivalent total scores on the HIM as a function of the cell weights, in effect having rating errors cancel each other out, it was necessary to assess the reliability of interview ratings in three ways:

1. by assessing interrater reliability as computed from the total scores generated in the interview using the HIM cell weight system

2. by a comparison of the percentage of responses recorded in the right, or "member" half of the matrix, and

3. by an analysis of the percentage of responses recorded in the lower-right hand, or "member-work", quadrant of the matrix.

In all three cases, interrater reliability coefficients were .90+, indicating a high degree of internal consistency among raters.

There are some real or potential problems evident in the use of the HIM beyond the training of raters:

1. Hill does not provide an operational definition of a response unit, either by time or by numbers of words. Consequently, the practice of rating each cell change and each speaker change, statement by statement, is the method in general use. This could result in some confusion between raters. This also necessitates the use of highly trained raters if reliability is to result. One method used to overcome this difficulty is to have typescripts of the interviews made and to allow the raters to work from the typescripts. This latter procedure is cumbersome and unnecessary with well trained raters.
2. Some words or sentence fragments are hard to record as individual responses and so are included with previous or following statements. Terrill and Terrill, using the Leary circumplex, reported the same difficulty.[15] They found it necessary to add additional response categories to account for this contingency. While two of these additional four categories are already in the HIM, one - Speech Lacking Information - is not included and forms the class of speech which tends to be included in prior or following statements.

Unpublished communications from Hill indicate that he is adding three additional categories to his matrix, none of which apparently carry any known therapeutic value, in an attempt to alleviate this problem. He has called these categories X, U, O. X is defined as no content statements, U as unfinished statements, and O as no

meaning to rater, i.e., non-verbal or in-group jokes.[6]

Terrill and Terrill listed clarity of category definition as one of their major problems. This seems to be true of all measures of this type. The work of Hill, however, has resulted in a scoring manual which reduces such problems to a minimum.[9]

3. While objectivity is added to the scale by rating primarily verbal content, affect is not always accounted for on the HIM.

4. Burks attempted to add additional information by developing a scale on which non-verbal behavior could be recorded.[3] Such behavior is an unknown with the HIM. The use of this Burks scale, however, necessitates the video taping of interviews, a procedure requiring equipment not always readily available.

5. The statement by statement rating system costs the researcher much information in terms of interview progression and communications source. This shortcoming is hard to overcome in group analysis due to the number of individuals who speak. In dyad research this is not a particularly difficult problem. By identifying each therapist or counselor statement with the letter T and each client statement with the letter C the rater can record the source of all communications. The addition of a consecutive numerical subscript to each therapist and counselor statement permits the charting of the entire interview flow. Additional information such as which participant led the interview toward, or tried to avoid, the more highly therapeutic interactions is obtained by this procedure.

Circling inappropriate responses, that is responses not logically atuned to the preceding statement also adds information for the researcher. It provides information relative to the counselor's skills and the client's willingness to discuss threatening topics. Attention to this interview facet can be related to the attending research of Ivey and his colleagues.[10] The resultant of these procedures provides interview data which can not only be analyzed in the standard way utilizing given cell weights but also provides a graphic description of the entire interview.

Another possible modification of the rating procedure is to have the raters list the statements in sequence by cell designation and adopt the results to the matrix after rating is completed.

6. One last problem remains. How much of any one interview should be rated? Obviously the entire interview may be rated. Alternatives to this may be the rating of short segments from various parts of the interview or some longer segment hypothesized to be that most productive for client change or of specific interest to the researcher. The taking of short time segments throughout the interview seems to be the procedure most often used.

In conclusion, this report has attempted to provide four types of data: (1) that relative to the measurement characteristics of the HIM, (2) the uses which have been made of it in individual counseling research, (3) problems involving its use, and (4) suggested extensions of the current scoring procedures.

As a measurement device the HIM appears to fulfill the conditions necessary for instruments of its type better than any of the other scales currently available.

References

1. Anderson, A. Personal communication. University of Minnesota. July, 1969.

2. Boyd, R. E. Counselor interests as a factor in counselor effectiveness. Unpublished doctoral dissertation, University of Minnesota, 1969.

3. Burks, H. M. The development and evaluation of an instrument for measuring selected dimensions of school counselor behavior in the interview. Unpublished doctoral dissertation, University of Minnesota, 1968.

4. Cannon, J. R. Effects of rater level of functioning and experience on the ability to discriminate counselor's levels of functioning. Unpublished doctoral dissertation, State University of New York at Buffalo, 1968.

5. Freedman, M. B., Leary, T. F., Ossorio, A. G., and Coffey, H. S. The interpersonal dimension of personality. Journal of Personality, 1951, 20, 143-161.

6. Gibson, D. Personal communication. University of Minnesota. March, 1970.

7. Hellervik, L. W. An operant conditioning approach to changing counselor interview behavior. Unpublished doctoral dissertation, University of Minnesota, 1968.

8. Hill, W. F. Hill interaction matrix. Los Angeles: University of Southern California, 1965.

9. Hill, W. F. Hill interaction matrix scoring manual. Los Angeles, University of Southern California, 1965.

10. Ivey, A. E., Normington, C. J., Miller, C. D., Morrill, W. H., and Haase, R. F. Microcounseling and attending behavior. Journal of Counseling Psychology Monograph Supplement, Part 2, 1968.

11. Lee, G. The effects of social reinforcement and some other experimental treatments on the within-interview verbal behavior of counselors. Unpublished doctoral dissertation, University of Minnesota, 1968.

12. Peterson, D. R. The clinical study of social behavior. New York: Appleton-Century-Crofts, 1968.

13. Seligman, M. and Sterne, D. M. Verbal behavior in therapist-led leaderless and alternating groups psychotherapy sessions. Journal of Counseling Psychology, 1969, 16, 325-328.

14. Stollak, G. E., Guerney, B., and Rothberg, M. Psychotherapy research. Chicago: Rand McNally, 1966.

15. Terrill, J. M. and Terrill, R. E. A method for studying family communications. Family Process, 1965, 4, 259-290.

16. Truax, C. B. and Carkhuff, R. R. Toward effective counseling and psychotherapy. Chicago: Aldine Publishing Company, 1967.

17. Vingoe, F. J. and Antonoff, S. R. Personality characteristics of good judges of others. Journal of Counseling Psychology, 1968, 15, 91-93.

18. Wellman, F. E. A conceptual framework for the derivation of guidance objectives and outcome criteria: preliminary statement in Whitely, John M. Research in counseling. Columbus: Charles Merrill, 1967.